

Are randomized trials the best way to judge the efficacy of surgical procedures?

Timothy J. Gardner, MD

Thoracic surgery is nearing its first century anniversary as a unique medical specialty. In addition, the year 2010 marks the 93rd anniversary of the founding of the American Association for Thoracic Surgery. What a century this has been for thoracic surgery! We have witnessed previously unimaginable accomplishments in the surgical treatment of diseases of the lung, esophagus, and heart. In the early part of the last century, surgery for infectious complications of the chest was common. Lung surgery for infection and cancer was next to develop as controlled ventilation was mastered. Cardiac surgery evolved dramatically from relatively simple “closed chest” procedures for congenital heart defects in the middle years of the last century. After the refinement of cardiopulmonary bypass 50 years ago, the range and numbers of cardiac surgical procedures exploded. As we reflect on the current success of our specialty with its broad scope of formidable surgical capabilities, we should challenge ourselves to ask to what extent thoracic surgery has matured as a scientific endeavor and a highly successful clinical specialty. Some have questioned the specialty’s commitment to pursuing evidence-based surgical care. Are we committed only to those surgical procedures that can be scientifically validated to be effective? It is in this context we should pose the question of whether “randomized controlled clinical trials are the best way to judge the efficacy of surgical operations.” There is an important stipulation to make, however, when discussing the importance of randomized trials to determine the appropriateness of any surgical procedure. Despite a commitment to the practice of evidence-based medicine and operating only when we are confident about the long-term consequences of the surgery, we must acknowledge the critical role of surgical innovation, technical development, and iterative refinements of surgical procedures. The remarkable heritage of our specialty has been the development of new, increasingly complex and highly effective surgical procedures. Over this past century, thoracic surgeons have developed an awesome array of operations that have treated and cured millions of patients. Along with these innovations and developments, however, have been the numerous operations that have been unsuccessful

and, in some instances, harmful to our patients. Ensuring that our surgical procedures are effective and necessary is the rationale and justification for conducting randomized surgical trials.

EVIDENCE-BASED SURGERY?

How is evidence of safety and efficacy generated for surgical operations? Consistent with the innovative history of surgical therapy, case reports have been the usual first-line method of dissemination of surgical results. Subsequent single-center or single-surgeon reports of case series add supportive but only limited evidence.¹ Although many landmark surgical innovations have been reported for the first time as case reports or small case series, these initial case series may be inadequate to demonstrate clear evidence of efficacy. Readily acknowledged limitations of such case series include patient selection bias, possible lack of general applicability when limited to procedures done by a single operator, and often an inadequate length and degree of follow-up assessment to generate confidence in the reported outcome. Reports by experienced and trusted surgeons create excitement and frequently generate additional case series and technical refinements by other surgeons. Although this is the usual pathway in the evolution of a new operation, the history of thoracic surgery is replete with case reports and surgical series of unsuccessful, unsafe, or ill-advised procedures. This trial-and-error approach may be appropriate and even necessary in the developmental stage of a new surgical procedure. When a new operation is reported to be safe and effective by several surgeons at more than one site, and the new procedure is being considered for wide adoption, there should be sufficient interest in the new operation to conduct an observational study that involves some standardization of the procedure and an attempt to involve multiple surgeons at a variety of centers. This important next step to provide supportive outcomes data to the anecdotal evidence of efficacy derived from case series is, unfortunately, rarely done. Even less common is the prospective, multicenter randomized trial that is widely acknowledged as the standard way to develop the appropriate evidence basis for any new therapy, including procedural therapies. Despite all of the formidable challenges and obstacles inherent in performing randomized surgical trials, such properly conducted studies are and will remain the gold standard for judging safety and therapeutic efficacy.

As specific surgical procedures are widely adopted and offered to patients, there is the understandable expectation by those patients, their referring physicians, and the public that

From the Christiana Care Health System, Center for Heart and Vascular Health, Christiana Hospital, Newark, Del.

Address for reprints: Timothy J. Gardner, MD, Christiana Care Health System, Center for Heart and Vascular Health, Christiana Hospital, Management Suite 1218, PO Box 6001, Newark, DE 19718 (E-mail: tgardner@christianacare.org).

J Thorac Cardiovasc Surg 2010;140:739-42

0022-5223/\$36.00

Copyright © 2010 by The American Association for Thoracic Surgery

doi:10.1016/j.jtcvs.2010.05.036

the safety and efficacy of the operation have been well established and can be performed safely by multiple surgeons at many different sites. Desperately ill patients may be willing to submit to a new surgical procedure when the operation is offered as a last resort. A surgeon may be willing to undertake an innovative procedure for an urgent indication when he or she is reasonably confident of his or her surgical and clinical abilities to safely perform the operation. It goes without saying, however, that surgeons have the broader responsibility to understand the full implications and consequences of the procedures they are performing. In addition, despite the traditional confidence in which surgeons are held by other physicians and prospective patients, public reporting of surgical outcomes is increasingly expected, if not demanded. Asking surgeons for clear evidence of efficacy can be expected to be raised more often in the future, not just by individual patients but by insurers, other providers, and public health policy makers. In the context of health care system reform in the United States, the use of comparative effectiveness of medical therapy is being discussed broadly and has become a mandate in the recent federal health care legislation. Although health policy experts stress that comparative effectiveness is a broader concept than simple "cost-effectiveness," major invasive and expensive therapies will be challenged for appropriateness on the basis of results and costs. Furthermore, the evolution of less invasive therapies that compete with or replace established and traditional surgical procedures, including percutaneous and endovascular interventions being performed by non-thoracic surgeons, will mandate that we participate in comparative effectiveness analyses. Well-conducted randomized trials will go a long way to establishing the evidence basis for the effectiveness of many commonly performed and important thoracic surgical operations.

EQUIPOISE: WHAT IS IT?

A common definition of equipoise when used in the context of medical therapy is as follows: Equipoise defines a state of uncertainty regarding the comparative merits of 2 different treatment options. It is little wonder that few surgeons are familiar with this terminology that was introduced in the context of clinical trials by Benjamin Freedman, a medical ethicist from McGill University, in the *New England Journal of Medicine* in 1987.² Freedman's article is titled "Equipoise and the Ethics of Clinical Research." He argues that each clinical trial should begin with an honest null hypothesis, indicating that the clinical investigator is in a state of genuine uncertainty regarding the merits of a new treatment compared with an established or accepted treatment. Freedman contends that a physician who knows that one treatment is superior to another cannot recommend participation in a comparative trial. On the other hand, he describes the concept of "clinical equipoise," which represents genuine uncertainty within the expert medical community, allowing for the indi-

vidual who may have some degree of bias about one treatment versus the other to ethically participate in the trial.

Surgeons rarely admit to being in doubt about a proposed operation. Uncertainty about a proposed operation is not a comfortable position for surgeons. Patients, especially those who are about to undergo a major operation with short-term mortality risk, do not expect to detect uncertainty about the effectiveness or safety of an operation from their surgeon. It is one thing to be told that a new drug that has been shown to be safe in preliminary trials may afford added therapeutic benefit compared with an established medication; it is more challenging to offer a patient randomization in a surgical trial with nothing more than a possibility that a new surgical procedure may provide some potential benefits compared with the established surgical approach. Patients, families, and referring physicians often become uncomfortable when their surgeon suggests that he or she is uncertain about which course of treatment to recommend. On the other hand, skilled and committed clinical investigators have been highly successful in reassuring anxious patients that comparative clinical studies, especially trials investigating new options for cancer therapies, may afford the possibility of even more successful treatment than that afforded by the established therapy.

This dilemma that is created by the surgeon appearing to be uncertain and indecisive demonstrates how challenging it is for many surgeons to admit that there is uncertainty about the effectiveness of our procedures. We are expected to be knowledgeable and decisive. Our work demands that we be certain about our skills. We are expected to function in an authoritative manner. After all, we are the captain of the ship, and even others on the surgical team are uncomfortable when we are or appear indecisive. Equipoise seems to be an oxymoron for an established surgeon. We would not be undertaking a major operation with a mortality risk for the patient unless we were certain of what we were proposing to do.

Finally, there is the challenge related to the pride of ownership for many surgeons, in particular, those innovators who have developed and championed new procedures for treating a disease or life-threatening condition. Surgeons with proprietary interests in devices are understandably conflicted. Those surgeons have worked hard to master a difficult procedure and may be resistant to consider engaging in a trial comparing a safe and comfortable technique with another potentially better operation. These are just some of the explanations for the unwillingness of some surgeons to achieve the state of equipoise that is required to convincingly participate in truly randomized clinical trials. Thoracic surgeons as a specialty group have not demonstrated much inclination to embrace the notion of equipoise. We are taught to be thorough and scientific in our evaluation of clinical data and decisive in our choice of therapeutic options. Expressions of doubt, inconsistency, and indecision are clinical

traits that are not acceptable in residency programs, when being examined by certifying boards, or rewarded with additional patient referrals when interacting with physicians. The irony of this situation is that the demand for targeted and quick decisions, even when meant to be based on evidence, often results in poorly informed clinical actions, including the choice of surgical procedures, the consequences of which are poorly understood. Determining the evidence base for our clinical and surgical decisions can be effectively accomplished by the appropriate use of well-designed and executed randomized clinical trials.

WHAT CHARACTERIZES A WELL-DESIGNED AND EXECUTED SURGICAL TRIAL?

The most common design type for surgical trials is the individually randomized, 2-group, parallel trial. The term “controlled” is redundant in most instances and is used simply to indicate that a control treatment group that receives no treatment or standard treatment is being compared with a new treatment. Trials that compare a major surgical operation with a non-operative control arm, although important and quite useful, have been challenging in terms of recruitment. Even when the surgical procedure under consideration holds substantial promise of benefit based on preliminary case reports, it can be difficult to convince patients to submit to randomization into a surgical treatment arm versus medical treatment alone, when the operation has mortality and major morbidity risks. Nonetheless, there have been important and successful randomized surgical trials, such as the Coronary Artery Surgery Study³ and the National Emphysema Treatment Trial,⁴ that have compared a surgical treatment patient group with a control group that received medical treatment only. It has been generally easier to enroll patients into randomized surgical trials when patients in both parallel arms of the study are undergoing an operative procedure. Two recent examples of large multicenter randomized surgical trials comparing different operative procedures are the VA Randomized On/Off Bypass Trial⁵ and the Coronary Bypass Surgery with or without Surgical Ventricular Reconstruction Trial.⁶

All of these trials share important characteristics of well-conducted randomized trials. Many surgeons at multiple sites participated, creating a “real-world” context and mitigating the influence of surgeons with exceptional technical experience or proficiency. The requirement for adequate statistical power was met in that the trials had adequate projected sample sizes and actual patient enrollment to achieve statistical confidence that the outcomes were not a result of chance. Inclusion and exclusion criteria were well defined, ensuring that patients in each treatment arm were comparable. Strictly enforced allocation into each arm resulted in true randomization and eliminated the possibility of selection bias. Patients were followed according to the initial treatment allocation regardless of whether the patient

actually received the allocated intervention, again mitigating the influence of post-randomization selection bias. Predefined primary and secondary outcomes were assessed. Both early and late outcomes are observed. It is noteworthy that only late outcomes assessment in the Bypass Angioplasty Revascularization Investigation Trial that compared percutaneous coronary balloon angioplasty with coronary artery bypass grafting demonstrated a difference in survival among diabetic patients, favoring those who had undergone surgery.⁷

An important aspect of a well-executed clinical trial includes appropriate documentation and reporting of the trial in a biomedical journal. Editors of many medical journals, but few surgical journals, have adopted reporting standards using the Consolidated Standards for Reporting of Trials checklist to ensure inclusion of all critical information required for interpretation of trial results.⁸ Compliance with these standards confers greater credibility to the reported trial and will facilitate use of the trial in future high-quality meta-analyses. Another useful addition to a well-conducted clinical trial is a registry of screened or eligible patients. Because a relatively small percentage of screened patients are actually enrolled in a randomized trial, it can be informative to assess the reasons for ineligibility or exclusion from the trial.

MOVING FORWARD IN DEFENSE OF RANDOMIZED SURGICAL TRIALS

As most surgeons readily acknowledge, and as commented on above, randomized surgical trials are challenging to conduct. In some ways, randomization of a surgical procedure goes against some of the most respected traditions of our specialty. We are taught from our earliest days as surgical residents to determine specific indications for specific operations and to decisively choose the right operation for a patient in need of surgery. We continue to look to senior surgeons and experts for guidance about what operation is most appropriate for which patient. Our respect for this surgical authority is deeply embedded in our thinking, as is evident from our dependence on presentations at specialty meetings and publications in medical journals for determining our choice of operations and treatment plans. The individual surgeon’s technical skills and confidence also influence his or her choice of an appropriate operation. Asking a surgeon to participate in a randomized trial in which an unfamiliar procedure may be mandated is not comfortable for the surgeon or safe for the patient, nor would participation by that surgeon be ethical. The understandable bias that derives from one’s own surgical experience or has resulted from the influence of respected colleagues challenges the degree of equipoise that is required of a surgical investigator. Equally formidable are the practical obstacles to performing clinical trials today, including the lack of available funding, absence of clinical research infrastructure in many surgical programs,

and human subject research regulations that have become increasingly burdensome.

On the other hand, the medical profession and, in particular, surgeons are being challenged to validate our proposed medical treatments and procedures. We will be increasingly required to demonstrate and document the safety and effectiveness of our surgical treatments. Randomized trials are at the peak of the evidence tree and are the best way to provide that expected evidence. Thoracic surgery's own Practice Guidelines recognize that data derived from multiple randomized clinical trials allow us to make Level A treatment recommendations. Although Level B evidence is obtained from a single randomized trial or from nonrandomized reports, most choices of treatment and surgical options today are only at evidence Level C, that is, evidence that is based on consensus expert opinion. This is the same "evidence" that relies on surgeons' experiences and case reports. The likelihood, if not the prediction, of erroneous treatment choices that are based on such expert opinion is substantial. Today's expert is tomorrow's fool dunced, duped not from a lack of effort or good intent, but from bias, limited experience, misinterpretation, or lack of awareness or acknowledgement of other treatment options. As mature

and respected specialists, thoracic surgeons must accept the challenge of participation, whenever feasible, in well-designed randomized trials. Our credibility as clinical scientists demands our leadership in such efforts and broad participation by thoracic surgeons.

References

1. Treasure T. Are randomized trials needed in the era of rapidly evolving technologies? *Eur J Cardiothorac Surg*. 2009;35:474-9.
2. Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med*. 1987;317:141-5.
3. CASS Principal Investigators and their associates. Coronary artery surgery study (CASS): a randomized trial of coronary artery bypass surgery. Survival data. *Circulation*. 1983;68:939-50.
4. National emphysema treatment trial research group. A randomized trial comparing lung-volume-reduction surgery with medical therapy for severe emphysema. *N Engl J Med*. 2003;348:2059-73.
5. Shroyer AL, Grover FL, Hattler B, Collins JF, McDonald GO, Kozora E, et al. On-pump versus off-pump coronary-artery bypass surgery. *N Engl J Med*. 2009;361:1827-37.
6. Jones RH, Velasquez EJ, Michler RE, Sopko G, Oh JK, O'Connor CM, et al. Coronary bypass surgery with or without surgical ventricular reconstruction. *N Engl J Med*. 2009;360:1705-17.
7. The Bari investigators. The final 10-year follow-up results from the BARI randomized trial. *J Am Coll Cardiol*. 2007;49:1600-6.
8. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomized trials. *Ann Intern Med*. 2010;152:1-7.